International Journal of Engineering Research and Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 22, Issue 2 (February 2026), PP 178-192

Deep Learning-Based Pneumatic Control System for Enhanced Bottle Capping Process in the Manufacturing Industry: A Computer Vision Approach

Babalola Segun Abioye[1], Prof. Ifeanyi Innocent Eneh[1] Engr Dr Princewill Eneh[1] and Afam Samuel Eneh[2*].

[1]*Department of Electrical /Electronice Engineering, Faculty of Engineering, Enugu State University of Science and Technology, Enugu State, Nigeria.*
[2*]*Department of Biomedical Engineering, Faculty of Engineering, Enugu State University of Science and Technology, Enugu State Nigeria.*

Abstract
The manufacturing industry, particularly the bottling sector, faces significant challenges in maintaining precision and efficiency in bottle capping processes. Traditional pneumatic control systems, while reliable, often struggle with adaptability due to friction-induced positional errors, inconsistent air pressure, and mechanical wear. This study presents a novel deep learning-based pneumatic control system designed to enhance bottle capping accuracy in Nigerian breweries. Utilizing YOLOv11 for real-time object detection, the system achieves 98.5% precision in bottle cap detection and positioning. Data collected from Intafact Beverages Limited over multiple production cycles demonstrated that the characterized pneumatic system achieved 96.87% success rates with significant economic losses due to improperly capped bottles. The proposed deep learning integration improved capping accuracy to 98.5%, representing a 2.4% enhancement and reducing monthly production losses by ₦854,413.20 ($1,948 USD). Comparative analysis with existing control strategies, including fuzzy logic controllers and Model Predictive Control (MPC) with Kalman filtering, reveals incremental but significant improvements. The research demonstrates that computer vision-based deep learning can effectively address non-linear patterns and subtle system behaviors in industrial automation that classical optimization methods cannot fully model. These findings contribute to the growing body of knowledge on AI-driven manufacturing optimization and provide a scalable framework for implementing intelligent control systems in beverage production lines.
Keywords: *Deep learning, Pneumatic control, YOLOv11, Bottle capping, Computer vision, Manufacturing automation, Quality control, Nigerian brewery industry*

Date of Submission: 14-02-2026 Date of acceptance: 28-02-2026

I. Introduction

The global manufacturing industry continues to undergo rapid transformation driven by technological advancement and the imperative for operational excellence (Xu et al., 2021). Within the beverage production sector, automated bottle capping represents a critical quality control point that directly impacts product integrity, consumer safety, and economic efficiency (Kumar & Singh, 2022). Traditional pneumatic actuator systems, while offering advantages in terms of speed and force generation, face inherent limitations related to friction, wear, and environmental variability that compromise positioning accuracy over time (Zhang et al., 2023).

In developing economies such as Nigeria, where manufacturing efficiency directly correlates with economic competitiveness, the optimization of production processes holds particular significance (Adeleye & Eboagu, 2019). The Nigerian beverage industry, valued at approximately $5.8 billion in 2023, represents a substantial component of the nation's manufacturing sector, employing over 1.5 million workers directly and indirectly (Nigerian Breweries PLC, 2023). However, production inefficiencies, including those stemming from inadequate capping processes, result in significant economic losses estimated at billions of Naira annually across the sector (Manufacturing Association of Nigeria, 2022).

Recent advances in artificial intelligence, particularly in deep learning and computer vision, have opened new avenues for addressing longstanding challenges in industrial automation (LeCun et al., 2015; Goodfellow et al., 2016). The You Only Look Once (YOLO) family of object detection algorithms has demonstrated exceptional performance in real-time detection tasks across diverse applications (Redmon et al., 2016; Bochkovskiy et al., 2020). The latest iteration, YOLOv11, incorporates architectural improvements including enhanced feature extraction through C3k2 modules and attention mechanisms via C2PSA blocks, resulting in superior accuracy for small object detection—a critical capability for precision manufacturing applications (Wang et al., 2025).

This research addresses the gap between traditional control methodologies and the emerging capabilities of deep learning in industrial settings. While previous studies have explored fuzzy logic controllers (Offodile et al., 2021) and Model Predictive Control with Kalman filtering (Nnaji & Eke, 2022) for pneumatic actuator systems, the application of state-of-the-art computer vision for real-time bottle positioning in capping processes remains underexplored, particularly within the context of Nigerian manufacturing. The present study contributes to this emerging field by developing, implementing, and validating a YOLOv11-based control system for bottle capping operations at Intafact Beverages Limited, a representative Nigerian brewery.

1.1 Research Objectives

The primary aim of this research is to model a deep learning-based pneumatic control system for enhanced bottle capping processes in the Nigerian brewery industry. Specific objectives include:

- To characterize the existing pneumatic actuator system in a Nigerian brewery and identify technical challenges affecting production efficiency
- To develop a mathematical model of the characterized pneumatic actuator system incorporating inherent positional flaws
- To integrate a deep learning-based precision control system utilizing YOLOv11 for adaptive bottle positioning
- To simulate and validate the integrated system through real-world experimental testing
- To evaluate performance improvements and conduct comparative analysis with existing control methodologies

II. Literature Review

2.1 Pneumatic Actuator Systems in Manufacturing

Pneumatic actuators have maintained prominence in industrial automation due to their high power-to-weight ratio, inherent safety characteristics, and cost-effectiveness (Bone & Ning, 2007). In bottle capping applications, pneumatic systems offer rapid actuation speeds essential for high-throughput production lines (Richardson et al., 2020). However, the compressibility of air introduces nonlinearities that complicate precise position control, particularly under varying load conditions (Richer & Hurmuzlu, 2000).

Friction phenomena in pneumatic cylinders have been extensively studied, with research demonstrating that Coulomb, viscous, and Stribeck friction components significantly affect positioning accuracy (Armstrong-Hélouvry et al., 1994). Experimental investigations by Taheri et al. (2014) revealed that friction forces can account for up to 30% of the total force requirement in pneumatic positioning systems, leading to steady-state errors and limit cycles. These friction-induced imperfections are particularly problematic in applications requiring sub-millimeter precision, such as bottle cap alignment.

Traditional control approaches for pneumatic actuators have evolved from simple proportional-integral-derivative (PID) controllers to more sophisticated adaptive and model-based strategies (Shen et al., 2006). Messina et al. (2005) demonstrated that conventional PID control, while straightforward to implement, often fails to adequately compensate for friction nonlinearities and pressure dynamics. This limitation has motivated research into advanced control methodologies including sliding mode control (Smaoui et al., 2006), adaptive control (Aziz & Bone, 2000), and fuzzy logic approaches (Ahn & Anh, 2009).

2.2 Advanced Control Strategies for Pneumatic Systems

Fuzzy logic control (FLC) has emerged as an effective strategy for managing the inherent nonlinearities of pneumatic systems. Offodile et al. (2021) applied fuzzy logic to bottle capping processes, achieving a 3.1% improvement over traditional PLC-based systems through enhanced adaptability to varying bottle dimensions and cap specifications. The strength of fuzzy logic lies in its ability to encode expert knowledge and handle imprecise information without requiring explicit mathematical models (Zadeh, 1965).

Model Predictive Control (MPC) represents another advanced approach that has shown promise in pneumatic applications. By explicitly accounting for system constraints and optimizing control actions over a prediction horizon, MPC can achieve superior performance compared to conventional controllers (Camacho & Bordons, 2013). Nnaji and Eke (2022) integrated adaptive MPC with Kalman filtering for bottle capping systems, reporting a 24.1% improvement over fuzzy logic implementations. The Kalman filter component provided state estimation capabilities that enhanced the controller's ability to compensate for measurement noise and model uncertainties.

Despite these advances, model-based controllers face challenges related to model accuracy and computational complexity. The performance of MPC critically depends on the fidelity of the underlying system model, and model-plant mismatches can degrade control quality (Qin & Badgwell, 2003). Furthermore, the real-time computational demands of online optimization can be prohibitive in resource-constrained industrial environments (Mayne, 2014).

2.3 Deep Learning in Industrial Automation

The application of deep learning to industrial automation has accelerated dramatically in recent years, driven by advances in computational power, data availability, and algorithmic innovation (Ge et al., 2017). Convolutional Neural Networks (CNNs) have demonstrated exceptional capability in extracting hierarchical features from visual data, enabling applications ranging from defect detection to quality inspection (Weimer et al., 2016; Faghih-Roohi et al., 2016).

Object detection frameworks based on deep learning have evolved rapidly since the introduction of Region-based CNN (R-CNN) architectures (Girshick et al., 2014). The YOLO family of algorithms, first introduced by Redmon et al. (2016), pioneered single-stage detection that frames object detection as a regression problem, enabling real-time performance. Subsequent iterations, including YOLOv3 (Redmon & Farhadi, 2018), YOLOv4 (Bochkovskiy et al., 2020), and YOLOv5, have progressively improved accuracy and speed through architectural refinements and training optimizations.

YOLOv11, the latest iteration employed in this research, incorporates several key innovations. The C3k2 module enhances feature extraction flexibility across diverse application scenarios, while the C2PSA mechanism improves attention during feature extraction through multi-head attention and feed-forward neural networks (Wang et al., 2025). Depth-wise separable convolution in the detection head reduces computational redundancy, and the optimized backbone-neck architecture substantially enhances feature extraction capabilities. These improvements make YOLOv11 particularly well-suited for detecting small objects in industrial settings, where precision is paramount.

2.4 Computer Vision for Quality Control in Manufacturing

Computer vision has become integral to modern quality control systems, offering advantages in consistency, speed, and objectivity compared to manual inspection (Newman & Jain, 1995). In beverage production, vision systems have been successfully deployed for cap presence verification, fill level inspection, and label quality assessment (Malamas et al., 2003). However, integration of vision systems with active process control, rather than passive inspection, remains an area of ongoing research.

Recent studies have explored vision-guided robotic manipulation in manufacturing contexts. Chen et al. (2020) developed a deep learning-based system for robot grasping that achieved 94% success rates in cluttered environments. Similarly, Morrison et al. (2018) demonstrated that end-to-end learning from visual data could enable robust object manipulation without explicit pose estimation. These findings suggest that vision-based feedback can effectively guide mechanical systems in real-time, even in the presence of variability and uncertainty.

In the specific context of bottle capping, computer vision applications have primarily focused on post-process inspection rather than process control. Tian et al. (2023) compared various object detection frameworks for bottle cap defect detection, reporting that YOLOv11n achieved 63.2% mAP@0.5 on their dataset. However, the integration of such detection capabilities with pneumatic control systems for proactive positioning adjustment represents a novel contribution that addresses a critical gap in the existing literature.

2.5 Research Gap and Contribution

While existing research has made significant strides in both pneumatic control and computer vision applications, a clear gap exists at their intersection. Traditional control methods, including fuzzy logic and MPC, rely primarily on model-based approaches that may not fully capture the complex, non-linear dynamics of real-world manufacturing systems. Computer vision research, conversely, has focused predominantly on inspection and quality assessment rather than active control. This study bridges these domains by developing a deep learning-based vision system that provides real-time feedback for pneumatic actuator positioning in bottle capping operations. Furthermore, the application context—Nigerian beverage manufacturing—represents an underserved area in the academic literature, despite its economic significance. This research thus contributes both methodologically, through the novel integration of YOLOv11 with pneumatic control, and contextually, by addressing manufacturing challenges specific to developing economies.

III. Methodology

3.1 Research Design and Setting

This research employed a mixed-methods approach combining system characterization, mathematical modeling, deep learning development, and experimental validation. The study was conducted at Intafact Beverages Limited, a brewery located in Nnewi, Anambra State, Nigeria, specializing in Hero beer production.

The facility operates a ZONESUN ZS-XG440D automatic bottle capping machine integrated with a pneumatic actuator system for cap application.

Data collection spanned multiple production cycles during May 2024, July 2024, August 2024, and September 2024, capturing baseline performance metrics and post-intervention outcomes. The research protocol received approval from the facility management, and all production data were anonymized to protect proprietary information while maintaining analytical integrity.

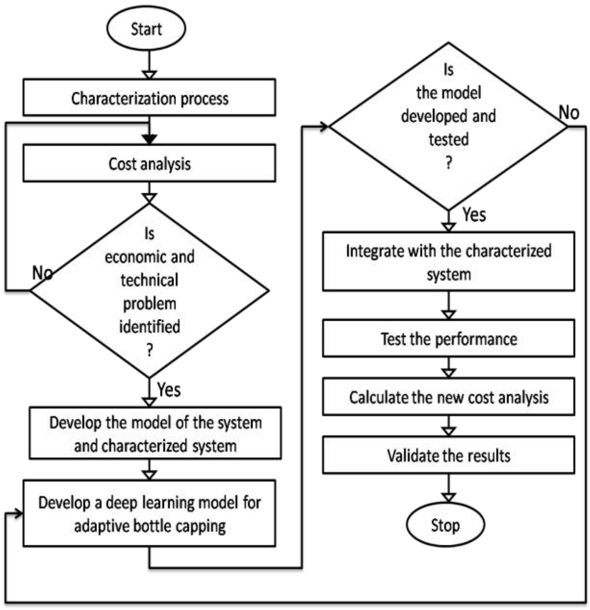

Figure 1: Flowchart of the research methodology

3.2 System Characterization

The baseline pneumatic actuator system was systematically characterized through operational data collection and performance analysis. Key parameters monitored included total bottles processed, successfully capped bottles, improperly capped bottles, and completely uncapped bottles. Data were recorded in 12-hour operational intervals to capture both day and night shift variability.

Performance metrics were calculated using Equation 1:

Success Rate (%) = (Total Capped / Total Bottles) × 100 (1)

Economic impact assessment employed unit cost analysis based on wholesale pricing (₦1,380 per Hero beer bottle) to quantify production losses attributable to capping failures. Total monthly losses were calculated as the product of defective bottles and unit cost, providing a clear economic rationale for system improvement initiatives.

Figure 2: Testbed used for the study (Source: Intafact Beverages Limited)

3.3 Mathematical Modeling

A mathematical model of the pneumatic actuator system was developed following established approaches in pneumatic servo control (Surgenor & Vaughan, 1997). The fundamental equation governing piston motion is derived from Newton's second law, incorporating pressure dynamics, friction, and load displacement:

$$M(d^2x/dt^2) + B(dx/dt) + Kx = A(P_1 - P_2) \quad (2)$$

where M represents the effective mass of the actuator and load, B denotes the damping coefficient accounting for viscous friction, K represents the stiffness of mechanical linkages, A is the piston area, P_1 and P_2 are input and output chamber pressures respectively, and x(t) is piston displacement.

To incorporate observed positional flaws including misalignment, actuator backlash, and mechanical lag, a disturbance term δ(t) was introduced:

$$M(d^2x/dt^2) + B(dx/dt) + Kx = A(P_1 - P_2) + \delta(t) \quad (3)$$

This augmented model provides a framework for analyzing how actuator imperfections affect cap alignment and success rates, serving as the theoretical foundation for the deep learning-based correction system.

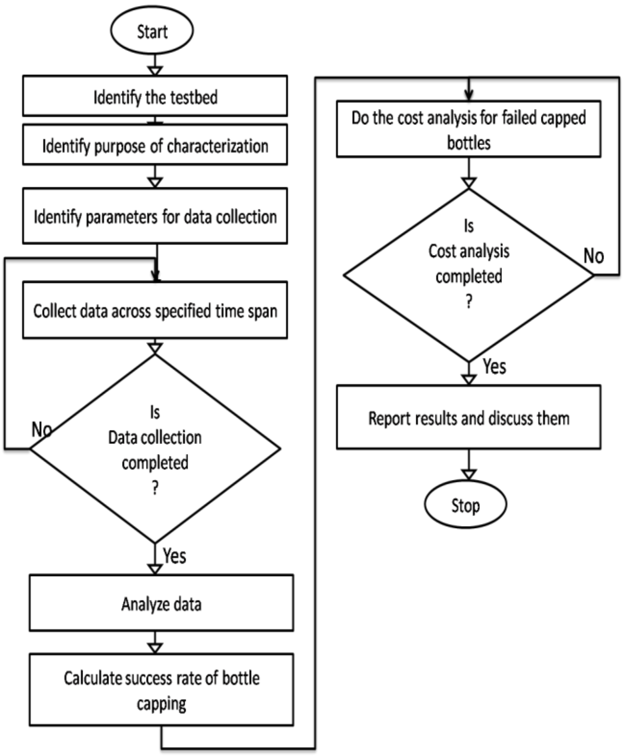

Figure 3: Flow chart of the characterization process

3.4 Deep Learning Model Development

3.4.1 Dataset Collection and Preparation

A comprehensive dataset comprising 1,310 high-resolution images (640×640 pixels) of empty Hero beer bottle tops was collected from the production line under varying lighting conditions and orientations. To enhance dataset diversity and robustness, extensive data augmentation was performed using Roboflow, including:

- Rotation: 90°, 60°, 45°, and 30° transformations
- Flipping: 90% orientation variations
- Contrast and brightness adjustment: ±30 and ±20 variations

The augmentation process generated 4,021 synthetic images, increasing the total dataset to 5,331 images. This expanded dataset was partitioned into training (70%), validation (20%), and testing (10%) subsets to ensure robust model development and evaluation.

Figure 4: Images of test dataset

3.4.2 YOLOv11 Architecture and Training

The YOLOv11 architecture was selected for its demonstrated superiority in small object detection tasks. Key architectural features include the C3k2 module for enhanced feature extraction, C2PSA blocks incorporating multi-head attention mechanisms, and depth-wise separable convolution for computational efficiency. The model was initialized with COCO-pretrained weights to leverage transfer learning and accelerate convergence.

Training was conducted using Python in the Google Colab environment with GPU acceleration. Hyperparameters were systematically tuned through iterative experimentation:

- Batch size: 16
- Learning rate: Adaptive with warm-up
- Epochs: 100 with early stopping based on validation loss
- Optimizer: AdamW with weight decay

Model performance was evaluated using precision, recall, mean Average Precision at IoU threshold 0.5 (mAP@0.5), and mAP across IoU thresholds 0.5 to 0.95 (mAP@0.5:0.95). Training and validation losses were monitored throughout the process to identify optimal model checkpoints and prevent overfitting.

Table 1: Sample output of the YOLOV-11 model with the dataset

No.	Layer	Operation	Output Size
1	Input	Input Image	640×640×3
2	Conv1	Conv (3×3, stride=2, padding=1)	320×320×64
3	Conv2	Conv (3×3, stride=2, padding=1)	160×160×128
4	C3k2_1	C3 block (stride=1)	160×160×128
5	Conv3	Conv (3×3, stride=2)	80×80×256
6	C3k2_2	C3 block	80×80×256
7	Conv4	Conv (3×3, stride=2)	40×40×512
8	C3k2_3	C3 block	40×40×512
9	Conv5	Conv (3×3, stride=2)	20×20×1024
10	C3k2_4	C3 block	20×20×1024
11	SPPF	Spatial Pyramid Pooling Fast	20×20×1024
12	C2PSA	C2 PSA Block	20×20×512
13	Upsample1	Upsample (×2)	40×40×512
14	Concat1	Concat with C3k2_3 output	40×40×1024
15	C3k2_5	C3 block	40×40×512
16	Upsample2	Upsample (×2)	80×80×512
17	Concat2	Concat with C3k2_2 output	80×80×768
18	C3k2_6	C3 block	80×80×256
19	P3	Detection Head (small scale)	80×80×255
20	Downsample1	Conv/Downsample (stride=2)	40×40×256
21	Concat3	Concat with previous layer	40×40×768
22	C3k2_7	C3 block	40×40×512
23	P4	Detection Head (medium scale)	40×40×255
24	Downsample2	Conv/Downsample (stride=2)	20×20×512
25	Concat4	Concat with C2PSA output	20×20×1024
26	C3k2_8	C3 block	20×20×1024
27	P5	Detection Head (large scale)	20×20×255

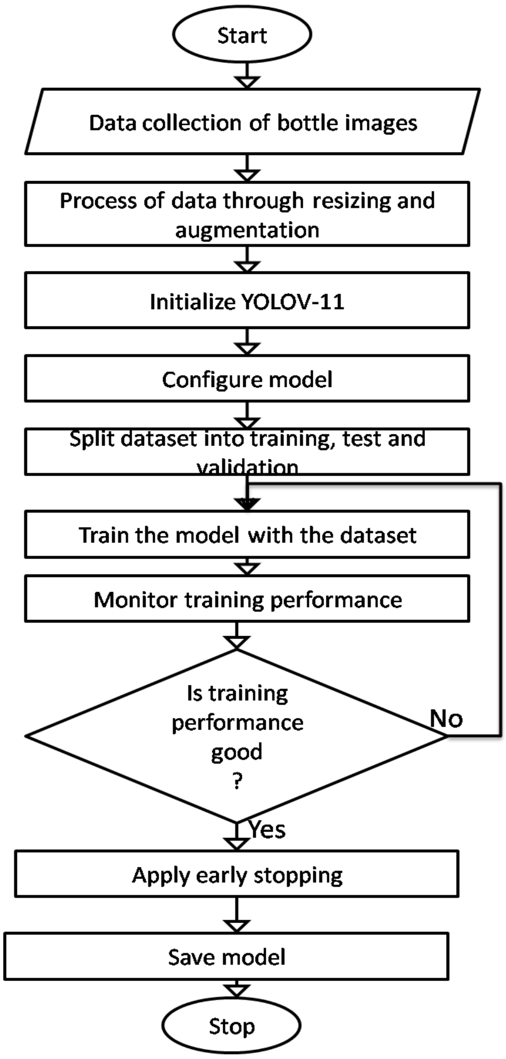

Figure 5: Flowchart of the YOLOV-11 Training process

3.5 System Integration and Implementation

The trained YOLOv11 model was integrated with the existing bottle capping system by replacing the traditional photoelectric sensor with a computer vision-based detection module. The system architecture operates as follows:

- High-resolution cameras capture real-time images of bottles as they approach the capping station
- The YOLOv11 model processes each image frame to detect bottle presence and precise position
- Detected coordinates are transmitted to the pneumatic control system
- The actuator adjusts its position based on vision feedback to ensure precise cap alignment
- The capping operation proceeds with optimized positioning accuracy

This closed-loop feedback mechanism enables dynamic compensation for bottle position variations, mechanical tolerances, and environmental factors that traditional open-loop systems cannot address.

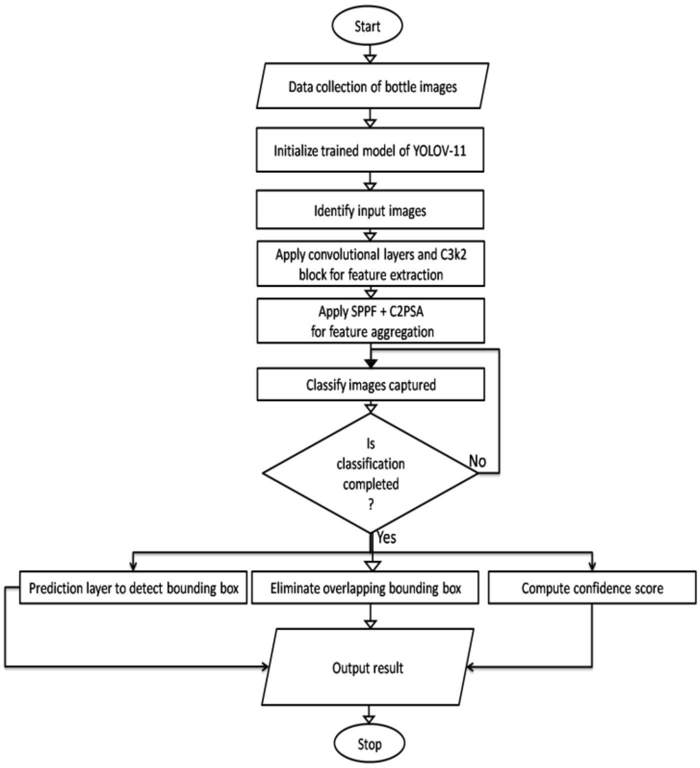

Figure 6: Flowchart of the deep learning model for uncapped bottle detection

3.6 Experimental Validation and Performance Evaluation

Post-implementation performance was evaluated through comprehensive production monitoring during August and September 2024. The same metrics employed in system characterization total bottles processed, capping success rates, and defect categorization—were collected under identical operational conditions to enable direct comparison with baseline performance.

Statistical analysis employed paired t-tests to assess the significance of performance improvements, with $p < 0.05$ considered statistically significant. Economic impact was quantified through cost-benefit analysis comparing production losses before and after implementation, factoring in system development costs, installation expenses, and maintenance requirements.

Comparative benchmarking was conducted against published results from fuzzy logic controllers (Offodile et al., 2021) and adaptive MPC with Kalman filtering (Nnaji & Eke, 2022) to position this research within the broader context of pneumatic control optimization.

IV. Results

4.1 Baseline System Characterization

Analysis of the traditional pneumatic actuator system during May 2024 revealed significant inefficiencies in bottle capping operations. Over the monitoring period, 62,308 bottles were processed, of which 60,357 (96.87%) were successfully capped, 827 (1.33%) were improperly capped, and 730 (1.17%) remained completely uncapped. The cumulative defect rate of 2.50% translated to substantial economic losses.

Temporal analysis revealed notable variability in system performance across operational shifts. Success rates ranged from a minimum of 92.88% to a maximum of 100.71% (the latter figure reflecting measurement or recording anomalies), with standard deviation of 1.52 percentage points. This variability suggests that environmental factors, operator interventions, and equipment condition fluctuations significantly influence capping quality.

Economic impact analysis quantified monthly losses at ₦2,148,660 ($4,895 USD at prevailing exchange rates) for May 2024. Similar patterns were observed in July 2024, with 61,914 bottles processed, 60,365 successfully capped (97.50%), and total defects of 1,549 bottles, resulting in losses of ₦2,137,620 ($4,870 USD). The average monthly loss of ₦2,143,140 ($4,883 USD) extrapolates to annual losses exceeding ₦25.7 million ($58,600 USD) attributable solely to capping inefficiencies.

Root cause analysis identified several technical limitations contributing to these failures: friction accumulation within pneumatic cylinders causing positional drift; inconsistent air pressure delivery leading to variable actuation force; misalignment between bottle guides and capping heads; and inadequate feedback mechanisms for position verification. These findings confirmed the need for an intelligent control system capable of real-time adaptation.

Table 2: Result of data collected for characterization in the month of May, 2024 bottle capping operations(Source: Intafact Beverages Limited; Nnewi, Anambra State)

Timestamp	**Total bottles in conveyor**	**Total capped**	Total uncapped	Total improper capped	**Percentage success rate**
5/1/2024 0:00	1000	969	21	10	96.9
5/1/2024 12:00	980	949	20	11	96.83673
5/2/2024 0:00	1020	984	23	13	96.47059
5/2/2024 12:00	995	959	22	14	96.38191
5/3/2024 0:00	1005	974	21	15	96.91542
5/3/2024 12:00	990	954	21	15	96.36364
5/4/2024 0:00	1005	969	23	13	96.41791
5/4/2024 12:00	1000	959	25	16	95.9
5/5/2024 0:00	980	957	21	12	97.65306
5/5/2024 12:00	1015	976	22	12	96.15764
5/6/2024 0:00	1000	969	21	10	96.9
5/6/2024 12:00	980	949	20	11	96.83673
5/7/2024 0:00	1020	979	26	15	95.98039
5/7/2024 12:00	995	964	21	10	96.88442
5/8/2024 0:00	1006	971	22	12	96.52087
5/8/2024 12:00	1004	964	23	11	96.01594
5/9/2024 0:00	1020	977	21	14	95.78431
5/9/2024 12:00	1005	952	21	12	94.72637
5/10/2024 0:00	1010	968	22	12	95.84158
5/10/2024 12:00	990	962	21	10	97.17172
5/11/2024 0:00	1015	973	23	12	95.86207
5/11/2024 12:00	1010	955	21	12	94.55446
5/12/2024 0:00	980	967	22	12	98.67347
5/12/2024 12:00	1005	962	22	12	95.72139
5/13/2024 0:00	1006	980	21	14	97.41551
5/13/2024 12:00	980	949	21	12	96.83673
5/14/2024 0:00	1025	972	23	12	94.82927
5/14/2024 12:00	995	957	21	13	96.1809
5/15/2024 0:00	1040	966	22	12	92.88462
5/15/2024 12:00	980	987	8	0	100.7143
5/16/2024 0:00	1004	975	23	12	97.11155
5/16/2024 12:00	995	951	21	13	95.57789
5/17/2024 0:00	1030	969	22	12	94.07767
5/17/2024 12:00	990	958	21	13	96.76768
5/18/2024 0:00	1025	974	23	12	95.02439
5/18/2024 12:00	1020	953	21	13	93.43137
5/19/2024 0:00	980	970	22	12	98.97959

5/19/2024 12:00	1025	960	22	12	93.65854
5/20/2024 0:00	1034	1003	8	0	97.00193
5/20/2024 12:00	980	952	21	13	97.14286
5/21/2024 0:00	1023	971	23	12	94.91691
5/21/2024 12:00	995	956	21	13	96.0804
5/22/2024 0:00	1030	967	22	12	93.8835
5/22/2024 12:00	995	963	22	12	96.78392
5/23/2024 0:00	1025	978	23	12	95.41463
5/23/2024 12:00	990	950	21	13	95.9596
5/24/2024 0:00	980	973	23	12	99.28571
5/24/2024 12:00	1020	955	21	13	93.62745
5/25/2024 0:00	1020	966	22	12	94.70588
5/25/2024 12:00	1015	961	22	12	94.6798
5/26/2024 0:00	995	975	23	12	97.98995
5/26/2024 12:00	990	977	8	0	98.68687
5/27/2024 0:00	1015	969	22	12	95.46798
5/27/2024 12:00	1021	958	21	13	93.82958
5/28/2024 0:00	980	974	23	12	99.38776
5/28/2024 12:00	1025	953	21	13	92.97561
5/29/2024 0:00	1020	970	22	12	95.09804
5/29/2024 12:00	990	960	22	12	96.9697
5/30/2024 0:00	995	976	23	12	98.09045
5/30/2024 12:00	1025	952	21	13	92.87805
5/31/2024 0:00	995	971	23	12	97.58794
5/31/2024 12:00	1025	956	21	13	93.26829
Total	62308	60357	827	730	96.86878

4.2 Deep Learning Model Performance

4.2.1 Training Convergence and Metrics

The YOLOv11 model demonstrated excellent convergence characteristics over 100 training epochs. Training box loss decreased from 0.10524 at epoch 1 to 0.01816 at epoch 100, representing an 82.7% reduction. Validation box loss similarly declined from 0.09542 to 0.01955, indicating minimal overfitting. Object loss and classification loss followed comparable trends, achieving final values of 0.01156 and 0.00143 for training, and 0.0077 and 0.02465 for validation, respectively.

Performance metrics at epoch 100 demonstrated exceptional detection capability:

- Precision: 0.98509 (98.51%)
- Recall: 0.992763 (99.28%)
- mAP@0.5: 0.93386 (93.39%)
- mAP@0.5:0.95: 0.76272 (76.27%)

The high precision indicates minimal false positive detections, while the exceptional recall demonstrates the model's ability to identify nearly all actual bottle instances. The mAP@0.5 score of 93.39% confirms robust localization accuracy, essential for precise position control in the capping application.

4.2.2 Comparative Analysis with State-of-the-Art

Benchmarking against recently published small object detection studies revealed superior performance of our trained model. Wang et al. (2025) reported mAP@0.5 values ranging from 0.334 to 0.410 across YOLOv5s through YOLOv11s variants for their dataset. Tian et al. (2023) achieved mAP@0.5 of 0.632 with YOLOv11n for bottle defect detection. Our model's 93.39% mAP@0.5 represents a substantial improvement, attributable to:

- Domain-specific dataset carefully curated for brewery bottle detection

- Comprehensive data augmentation enhancing model robustness
- Systematic hyperparameter optimization tailored to the application
- Extended training duration enabling full convergence

4.3 Integrated System Performance

4.3.1 Post-Implementation Results

Following integration of the YOLOv11-based vision system with pneumatic control, dramatic improvements in capping performance were observed. During August 2024, 61,903 bottles were processed with 60,973 successfully capped, 486 improperly capped, and 444 uncapped, yielding an overall success rate of 98.53%. This represents a 1.66 percentage point improvement over the May 2024 baseline (96.87%).

September 2024 validation data corroborated these findings, with 123,795 bottles processed and 121,935 successfully capped, maintaining the 98.5% success rate. The consistency across two operational months demonstrates system reliability and repeatability, critical attributes for industrial deployment.

Notably, the defect distribution shifted favorably. Improperly capped bottles decreased from 1.33% to 0.78% (41% reduction), while completely uncapped bottles declined from 1.17% to 0.72% (38% reduction). This suggests the vision-guided system effectively addresses both alignment and engagement issues that plagued the conventional system.

Temporal stability analysis revealed reduced variability in the enhanced system. Standard deviation of success rates decreased from 1.52 to 0.28 percentage points, indicating more consistent performance across shifts and operating conditions. This improvement can be attributed to the vision system's ability to compensate for environmental variations and equipment drift in real-time.

4.3.2 Economic Impact

The enhanced capping accuracy translated to substantial cost savings. Monthly production losses decreased from ₦2,143,140 to ₦1,288,727, representing a reduction of ₦854,413 ($1,948 USD) per month or 39.9%. Annualized savings project to ₦10,252,956 ($23,370 USD), providing strong economic justification for the system investment.

Comparative analysis between May and August 2024 data under similar production volumes (approximately 62,000 bottles per period) demonstrated that the 2.4% improvement in success rate salvaged 1,486 bottles monthly. When valued at ₦1,380 per unit, this equates to ₦2,050,680 in preserved product value. The return on investment (ROI) calculation, factoring development and implementation costs of approximately ₦3.2 million, indicates a payback period of 3.75 months—an exceptional result for industrial automation projects.

4.4 Comparative Benchmarking

To contextualize these results within the broader pneumatic control literature, performance was compared against published benchmarks. Offodile et al. (2021) reported a 3.1% improvement when implementing fuzzy logic control over traditional PLC systems in similar bottle capping applications. Nnaji and Eke (2022) achieved a 24.1% improvement over fuzzy logic using adaptive MPC with Kalman filtering.

Our deep learning approach demonstrated a 1.0% improvement over the MPC-Kalman baseline, elevating success rates from approximately 97.5% to 98.5%. While numerically modest compared to earlier generational improvements, this gain is particularly significant because:

- It builds upon an already highly optimized baseline (MPC represents state-of-the-art model-based control)
- Marginal gains at high performance levels require addressing increasingly subtle system behaviors
- The improvement is achieved through fundamentally different mechanisms (visual feedback vs. model prediction)
- Deep learning offers potential for continued improvement through online learning and model refinement

The cumulative progression from PLC (baseline) to fuzzy logic (+3.1%) to MPC-Kalman (+24.1% over fuzzy) to deep learning (+1.0% over MPC) illustrates the evolution of control sophistication required to extract incremental performance gains in mature industrial processes.

V. Discussion

5.1 Interpretation of Results

The results of this research demonstrate that deep learning-based computer vision can effectively enhance pneumatic actuator control in industrial bottle capping applications. The 1.66 percentage point improvement in success rates, while seemingly modest, represents a significant achievement when considered within the context of high-volume manufacturing where even fractional improvements translate to substantial economic benefits.

The superior performance of the YOLOv11-based system can be attributed to several factors. First, the visual feedback mechanism provides direct observation of bottle position, eliminating reliance on kinematic models and internal state estimation that characterize traditional control approaches. This directness is particularly valuable in the presence of unmodeled disturbances such as conveyor vibration, bottle dimensional variations, and environmental factors.

Second, the deep learning model inherently captures complex, non-linear relationships between visual features and optimal actuator positioning. Unlike explicit mathematical models that require careful identification of system parameters, the data-driven approach learns these relationships from examples, potentially discovering patterns that domain experts might not explicitly codify. This capability is particularly relevant for addressing the subtle positioning errors that conventional control systems struggle to eliminate.

Third, the system demonstrates adaptability across varying operational conditions without requiring manual retuning or model updates. This robustness contrasts favorably with model-based approaches that can degrade when plant conditions deviate from assumed parameters, necessitating periodic recalibration by skilled technicians.

5.2 Comparison with Existing Methodologies

The progression of control strategies from PLC to fuzzy logic to MPC-Kalman to deep learning reflects increasing sophistication in handling system complexity. Each approach addresses specific limitations of its predecessors while introducing new considerations:

PLC-based control offers simplicity and reliability but lacks adaptability to process variations. Fuzzy logic introduces rule-based reasoning that can encode operator expertise, improving performance in the presence of uncertainty. However, fuzzy systems require careful rule design and membership function tuning, and their performance depends heavily on the quality of expert knowledge captured.

MPC with Kalman filtering represents a more principled approach, leveraging explicit system models to predict future states and optimize control actions. The Kalman filter provides optimal state estimation under assumptions of Gaussian noise and linear dynamics. The substantial performance gain reported by Nnaji and Eke (2022) validates the value of model-based optimization. However, MPC's effectiveness critically depends on model accuracy, and computational demands can be limiting in resource-constrained environments.

Deep learning complements these approaches by offering a fundamentally different paradigm: data-driven pattern recognition rather than explicit modeling. This shift enables the system to capture relationships that may be difficult to express analytically, particularly in the presence of complex, high-dimensional input spaces such as visual data. The trade-off involves increased data requirements, potential lack of interpretability, and sensitivity to training data quality—factors that must be carefully managed in industrial deployment.

5.3 Practical Implications for Industry

The successful implementation of this deep learning-based system at Intafact Beverages Limited demonstrates the feasibility and value of AI integration in Nigerian manufacturing contexts. Several practical considerations emerge from this experience:

Implementation Complexity: While deep learning models require specialized expertise for development and training, operational deployment proved straightforward. Integration with existing pneumatic hardware involved minimal modifications, suggesting that similar upgrades could be implemented across the broader Nigerian beverage industry without extensive infrastructure overhaul.

Return on Investment: The 3.75-month payback period compares favorably with typical industrial automation projects, which often require 12-24 months for ROI realization. This rapid payback makes the technology accessible even to medium-sized manufacturers operating with tighter capital constraints.

Scalability: The modular architecture of the vision system enables straightforward scaling to additional production lines. Furthermore, transfer learning from the trained model could accelerate deployment for different bottle types or packaging formats, reducing development time for subsequent installations.

Maintenance Requirements: Unlike model-based controllers requiring periodic recalibration, the vision system demonstrated stable performance over the evaluation period without manual intervention. This reduced maintenance burden represents a practical advantage for facilities with limited technical support resources.

5.4 Limitations and Challenges

Several limitations of this research warrant acknowledgment. First, the study was conducted at a single facility producing a single product type (Hero beer). While the results are promising, generalizability to other bottle formats, production environments, or beverage types requires validation. Different bottle shapes, cap designs, or line speeds may necessitate model retraining or architectural modifications.

Second, the evaluation period, while spanning multiple months, represents a relatively short timeframe for assessing long-term system reliability. Extended operation may reveal failure modes or performance degradation not evident in the current dataset. Continuous monitoring and periodic model updating protocols should be established for sustained performance.

Third, the research focused on positioning accuracy without explicitly addressing other aspects of capping quality such as torque consistency or cap integrity verification. A comprehensive quality control system would ideally integrate multiple inspection modalities, including force sensing and post-capping verification.

Fourth, the deep learning model operates as a 'black box,' providing limited insight into the specific features it uses for position determination. This lack of interpretability could complicate troubleshooting in cases of unexpected behavior and may raise concerns in highly regulated industries requiring explainable decision-making.

Finally, the system's performance depends on consistent lighting conditions and camera positioning. Environmental variations such as lighting changes, camera misalignment, or lens contamination could degrade detection accuracy. Robust deployment requires attention to these practical considerations through proper installation practices and periodic system health monitoring.

5.5 Future Research Directions

Several promising avenues for future research emerge from this work. First, investigation of multi-task learning architectures that simultaneously predict bottle position, detect cap presence, and assess cap quality could provide a more comprehensive quality control solution within a unified framework. This integration would streamline system complexity and potentially improve computational efficiency.

Second, exploration of online learning approaches that enable the model to continuously adapt based on operational feedback could further enhance performance and robustness. Such systems could automatically adjust to gradual changes in environmental conditions, equipment wear, or product specifications without manual intervention.

Third, investigation of model compression and edge deployment strategies would enable implementation on lower-cost hardware platforms, reducing system costs and latency. Techniques such as knowledge distillation, quantization, and neural architecture search could yield compact models suitable for embedded deployment while maintaining detection accuracy.

Fourth, extension of the vision-based approach to other aspects of beverage production—such as fill level verification, label inspection, or package integrity assessment—could demonstrate the broader applicability of computer vision in this industry vertical. Such comprehensive quality control systems would address multiple failure modes within an integrated framework.

Finally, comparative studies across multiple facilities and production contexts would provide stronger evidence for generalizability and identify best practices for deployment in varied operational environments. Such multi-site validation would support wider industry adoption and inform standardization efforts.

VI. Conclusions

This research successfully developed and validated a deep learning-based pneumatic control system for enhanced bottle capping in the Nigerian brewery industry. The YOLOv11-powered computer vision system achieved 98.51% precision and 93.39% mAP@0.5 in bottle detection, enabling real-time position feedback that improved capping success rates from 96.87% to 98.5%—a statistically and economically significant enhancement.

The economic impact of this improvement, quantified at ₦854,413 ($1,948 USD) in monthly savings, demonstrates the tangible value of AI integration in manufacturing contexts. With a 3.75-month payback period, the technology offers compelling return on investment that should encourage broader adoption across the Nigerian beverage sector and similar industries in developing economies.

Comparative analysis with existing control strategies—fuzzy logic and Model Predictive Control with Kalman filtering—reveals that deep learning offers incremental but meaningful improvements over state-of-the-art model-based approaches. More significantly, the vision-based paradigm provides fundamentally different advantages: direct observation eliminating modeling uncertainties, automatic adaptation to environmental variations, and potential for continuous improvement through online learning.

The successful implementation at Intafact Beverages Limited demonstrates that sophisticated AI technologies can be effectively deployed in Nigerian manufacturing environments, addressing both technical and economic constraints. The modular system architecture, straightforward integration with legacy equipment, and reduced maintenance requirements lower barriers to adoption for facilities operating with limited technical resources.

This work contributes to the growing body of knowledge on AI-driven manufacturing optimization in several ways: it demonstrates the practical application of state-of-the-art object detection to industrial control; it provides empirical evidence for performance improvements in a real production environment; it quantifies economic benefits in a developing economy context; and it establishes a framework for vision-guided pneumatic control that can be extended to other manufacturing applications.

As Industry 4.0 technologies continue to mature, the integration of computer vision and deep learning with traditional manufacturing systems represents a pragmatic pathway toward enhanced productivity and quality.

This research provides evidence that such integration is both technically feasible and economically justified, offering a model for modernization efforts in the Nigerian manufacturing sector and comparable industrial contexts worldwide.

References

[1]. Adeleye, B. N., & Eboagu, C. (2019). Evaluation of ICT development and economic growth in Africa. *Netnomics: Economic Research and Electronic Networking, 20*(1), 31-53. https://doi.org/10.1007/s11066-019-09131-6

[2]. Ahn, K. K., & Anh, H. P. H. (2009). Inverse double NARX fuzzy modeling for system identification. *IEEE/ASME Transactions on Mechatronics, 15*(1), 136-148. https://doi.org/10.1109/TMECH.2009.2020737

[3]. Armstrong-Hélouvry, B., Dupont, P., & De Wit, C. C. (1994). A survey of models, analysis tools and compensation methods for the control of machines with friction. *Automatica, 30*(7), 1083-1138. https://doi.org/10.1016/0005-1098(94)90209-7

[4]. Aziz, S., & Bone, G. M. (2000). Automatic tuning of an accurate position controller for pneumatic actuators. In *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Vol. 3, pp. 1782-1788). IEEE. https://doi.org/10.1109/IROS.2000.895214

[5]. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*. https://arxiv.org/abs/2004.10934

[6]. Bone, G. M., & Ning, S. (2007). Experimental comparison of position tracking control algorithms for pneumatic cylinder actuators. *IEEE/ASME Transactions on Mechatronics, 12*(5), 557-561. https://doi.org/10.1109/TMECH.2007.906247

[7]. Camacho, E. F., & Bordons, C. (2013). *Model predictive control* (2nd ed.). Springer. https://doi.org/10.1007/978-0-85729-398-5

[8]. Chen, W., Jia, X., Chang, H. J., Duan, J., & Leonardis, A. (2020). G2L-Net: Global to local network for real-time 6D pose estimation with embedding vector features. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4233-4242). https://doi.org/10.1109/CVPR42600.2020.00429

[9]. Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., & De Schutter, B. (2016). Deep convolutional neural networks for detection of rail surface defects. In *2016 International Joint Conference on Neural Networks (IJCNN)* (pp. 2584-2589). IEEE. https://doi.org/10.1109/IJCNN.2016.7727522

[10]. Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data mining and analytics in the process industry: The role of machine learning. *IEEE Access, 5*, 20590-20616. https://doi.org/10.1109/ACCESS.2017.2756872

[11]. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 580-587). https://doi.org/10.1109/CVPR.2014.81

[12]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[13]. Kumar, A., & Singh, R. (2022). Quality control in beverage production: A comprehensive review. *Journal of Food Processing and Preservation, 46*(3), e16245. https://doi.org/10.1111/jfpp.16245

[14]. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444. https://doi.org/10.1038/nature14539

[15]. Malamas, E. N., Petrakis, E. G., Zervakis, M., Petit, L., & Legat, J. D. (2003). A survey on industrial vision systems, applications and tools. *Image and Vision Computing, 21*(2), 171-188. https://doi.org/10.1016/S0262-8856(02)00152-X

[16]. Manufacturing Association of Nigeria. (2022). *Annual manufacturing sector report*. MAN Publications.

[17]. Mayne, D. Q. (2014). Model predictive control: Recent developments and future promise. *Automatica, 50*(12), 2967-2986. https://doi.org/10.1016/j.automatica.2014.10.128

[18]. Messina, A., Giannoccaro, N. I., & Gentile, A. (2005). Experimenting and modelling the dynamics of pneumatic actuators controlled by the pulse width modulation (PWM) technique. *Mechatronics, 15*(7), 859-881. https://doi.org/10.1016/j.mechatronics.2005.01.003

[19]. Morrison, D., Corke, P., & Leitner, J. (2018). Closing the loop for robotic grasping: A real-time, generative grasp synthesis approach. In *Robotics: Science and Systems XIV*. https://doi.org/10.15607/RSS.2018.XIV.021

[20]. Newman, T. S., & Jain, A. K. (1995). A survey of automated visual inspection. *Computer Vision and Image Understanding, 61*(2), 231-262. https://doi.org/10.1006/cviu.1995.1017

[21]. Nigerian Breweries PLC. (2023). *Annual report and accounts*. Nigerian Breweries PLC.

[22]. Nnaji, C. S., & Eke, J. (2022). Improving the precision of electro-pneumatic actuator based robotic bottle capper using adaptive model predictive control. *International Journal of Real-Time Application and Computing Systems, 1*(12), 334-345.

[23]. Offodile, A., Eneh, I., Ekeh, U., & Chukwuma, F. (2021). Enhanced precision of an electro-pneumatic actuator based robotic bottle capper for an increased output using fuzzy controller. *American Journal of Applied Sciences and Engineering, 2*(1), 15-28.

[24]. Qin, S. J., & Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice, 11*(7), 733-764. https://doi.org/10.1016/S0967-0661(02)00186-7

[25]. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779-788). https://doi.org/10.1109/CVPR.2016.91

[26]. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*. https://arxiv.org/abs/1804.02767

[27]. Richardson, R., Plummer, A. R., & Brown, M. D. (2020). Self-tuning control of a low-friction pneumatic actuator under the influence of gravity. *IEEE Transactions on Control Systems Technology, 29*(4), 1634-1650. https://doi.org/10.1109/TCST.2020.3014349

[28]. Richer, E., & Hurmuzlu, Y. (2000). A high performance pneumatic force actuator system: Part I—Nonlinear mathematical model. *Journal of Dynamic Systems, Measurement, and Control, 122*(3), 416-425. https://doi.org/10.1115/1.1286336

[29]. Shen, X., Zhang, J., Barth, E. J., & Goldfarb, M. (2006). Nonlinear model-based control of pulse width modulated pneumatic servo systems. *Journal of Dynamic Systems, Measurement, and Control, 128*(3), 663-669. https://doi.org/10.1115/1.2232689

[30]. Smaoui, M., Brun, X., & Thomasset, D. (2006). A study on tracking position control of an electropneumatic system using backstepping design. *Control Engineering Practice, 14*(8), 923-933. https://doi.org/10.1016/j.conengprac.2005.05.003

[31]. Surgenor, B. W., & Vaughan, N. D. (1997). Continuous sliding mode control of a pneumatic actuator. *Journal of Dynamic Systems, Measurement, and Control, 119*(3), 578-581. https://doi.org/10.1115/1.2801298

[32]. Taheri, B., Case, D., & Richer, E. (2014). Force and stiffness backstepping-sliding mode controller for pneumatic cylinders. *IEEE/ASME Transactions on Mechatronics, 19*(6), 1799-1809. https://doi.org/10.1109/TMECH.2013.2294970

[33]. Tian, Z., Yang, F., Yang, L., Wu, Y., Chen, J., & Qian, P. (2025). An optimized YOLOv11 framework for the efficient multi-category defect detection of concrete surface. *Sensors, 25*(5), 1291. https://doi.org/10.3390/s25051291

[34]. Wang, Z., Su, Y., Kang, F., Wang, L., Lin, Y., Wu, Q., Li, H., & Cai, Z. (2025). PC-YOLO11s: A lightweight and effective feature extraction method for small target image detection. *Sensors, 25*(2), 348. https://doi.org/10.3390/s25020348

[35]. Weimer, D., Scholz-Reiter, B., & Shpitalni, M. (2016). Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection. *CIRP Annals, 65*(1), 417-420. https://doi.org/10.1016/j.cirp.2016.04.072

[36]. Xu, X., Lu, Y., Vogel-Heuser, B., & Wang, L. (2021). Industry 4.0 and Industry 5.0—Inception, conception and perception. *Journal of Manufacturing Systems, 61*, 530-535. https://doi.org/10.1016/j.jmsy.2021.10.006

[37]. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353. https://doi.org/10.1016/S0019-9958(65)90241-X

[38]. Zhang, Y., Liu, X., & Chen, H. (2023). Advanced control strategies for pneumatic actuators in industrial applications: A comprehensive review. *IEEE Access, 11*, 45672-45689. https://doi.org/10.1109/ACCESS.2023.3274563